Three.IV Matrix Operations

Linear Algebra

Jim Hefferon

http://joshua.smcvt.edu/linearalgebra

Sums and Scalar Products

Representing operations on linear functions

Recall the operations on linear functions of scalar multiplication and addition defined by: for $f, g\colon V \to W$, where $r \in \mathbb{R}$ the scalar multiple $r \cdot f$ function does this

$$\vec{v} \stackrel{r\cdot f}{\longmapsto} r \cdot (f(\vec{v}))$$

and the sum of the two functions does this.

$$\vec{v} \stackrel{f+g}{\longmapsto} f(\vec{v}) + g(\vec{v})$$

We will see how the matrix representations $\mathrm{Rep}_{B,D}(f)$ and $\mathrm{Rep}_{B,D}(g)$ combine to give $\mathrm{Rep}_{B,D}(r \cdot f)$ and $\mathrm{Rep}_{B,D}(f + g)$.

Representing operations on linear functions

Recall the operations on linear functions of scalar multiplication and addition defined by: for $f, g\colon V \to W$, where $r \in \mathbb{R}$ the scalar multiple $r \cdot f$ function does this

$$\vec{v} \xmapsto{r \cdot f} r \cdot (f(\vec{v}))$$

and the sum of the two functions does this.

$$\vec{v} \xmapsto{f+g} f(\vec{v}) + g(\vec{v})$$

We will see how the matrix representations $\mathrm{Rep}_{B,D}(f)$ and $\mathrm{Rep}_{B,D}(g)$ combine to give $\mathrm{Rep}_{B,D}(r \cdot f)$ and $\mathrm{Rep}_{B,D}(f + g)$.

Example Suppose that $f\colon V \to W$ is the linear map represented with respect to some bases B, D by this matrix.

$$F = \mathrm{Rep}_{B,D}(f) = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$$

We will find the matrix representing the function $6f$. That is, we will show how to produce the matrix representation of $6f$ from the matrix representation F of f.

Consider representations of vectors from the domain and codomain.

$$\mathrm{Rep}_B(\vec{v}) = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \quad \mathrm{Rep}_D(\vec{w}) = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

Consider representations of vectors from the domain and codomain.

$$\mathrm{Rep}_B(\vec{v}) = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \quad \mathrm{Rep}_D(\vec{w}) = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

Since $6\vec{w} = 6 \cdot (w_1\vec{\delta}_1 + w_2\vec{\delta}_2) = 6w_1\vec{\delta}_1 + 6w_2\vec{\delta}_2$, application of the function $6f\colon V \to W$ yields this representation of the codomain vector.

$$\mathrm{Rep}_D(6f(\vec{v})) = \mathrm{Rep}_D(6\vec{w}) = \begin{pmatrix} 6w_1 \\ 6w_2 \end{pmatrix}$$

So moving from the function f to the function $6f$ has the effect on the representations of vectors that it transforms the codomain vector by multiplying each entry by 6.

In this example the action of the function is represented as here.

$$\mathrm{Rep}_{B,D}(f) \cdot \mathrm{Rep}_{B}(\vec{v}) = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 2v_1 + v_2 \\ 3v_1 + 4v_2 \end{pmatrix}$$

In this example the action of the function is represented as here.

$$\mathrm{Rep}_{B,D}(f)\cdot\mathrm{Rep}_{B}(\vec{v})=\begin{pmatrix}2&1\\3&4\end{pmatrix}\begin{pmatrix}v_1\\v_2\end{pmatrix}=\begin{pmatrix}2v_1+v_2\\3v_1+4v_2\end{pmatrix}$$

We want the matrix that makes this true.

$$\mathrm{Rep}_{B,D}(6f)\cdot\begin{pmatrix}v_1\\v_2\end{pmatrix}=\begin{pmatrix}6(2v_1+v_2)\\6(3v_1+4v_2)\end{pmatrix}=\begin{pmatrix}12v_1+6v_2\\18v_1+24v_2\end{pmatrix}$$

Therefore, at least in this example

$$\mathrm{Rep}_{B,D}(4f)=\begin{pmatrix}12&6\\18&24\end{pmatrix}$$

so going from the function f to the function $6f$ has the effect on the representation F of multiplying all the entries by 6.

Example Suppose that $f, g\colon V \to W$ are linear maps represented with respect to some bases by these matrices.

$$F = \text{Rep}_{B,D}(f) = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \qquad G = \text{Rep}_{B,D}(g) = \begin{pmatrix} 5 & 8 \\ 7 & 6 \end{pmatrix}$$

We want the matrix representing the function $f + g$.

Example Suppose that $f, g\colon V \to W$ are linear maps represented with respect to some bases by these matrices.

$$F = \mathrm{Rep}_{B,D}(f) = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \qquad G = \mathrm{Rep}_{B,D}(g) = \begin{pmatrix} 5 & 8 \\ 7 & 6 \end{pmatrix}$$

We want the matrix representing the function $f + g$.

If $f(\vec{v}) = \vec{u}$ and $g(\vec{v}) = \vec{w}$ then $f + g$ does this.

$$\begin{aligned} \vec{v} \overset{f+g}{\longmapsto} \vec{u} + \vec{v} \\ &= u_1\vec{\delta}_1 + u_2\vec{\delta}_2 + w_1\vec{\delta}_1 + w_2\vec{\delta}_2 \\ &= (u_1 + w_1)\vec{\delta}_1 + (u_2 + w_2)\vec{\delta}_2 \end{aligned}$$

Thus for any $\vec{v}$, the effect on the representatives

$$\mathrm{Rep}_D(f(\vec{v})) = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad \mathrm{Rep}_D(g(\vec{v})) = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

of adding the functions is to add the column vectors.

$$\mathrm{Rep}_D((f + g)(\vec{v})) = \begin{pmatrix} u_1 + w_1 \\ u_2 + w_2 \end{pmatrix}$$

The particular functions in this example have this action on the representations.

$$\mathrm{Rep}_D(f(\vec{v})) = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 2v_1 + v_2 \\ 3v_1 + 4v_2 \end{pmatrix}$$

$$\mathrm{Rep}_D(g(\vec{v})) = \begin{pmatrix} 5 & 8 \\ 7 & 6 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 5v_1 + 8v_2 \\ 7v_1 + 6v_2 \end{pmatrix}$$

So we want the matrix with this effect.

$$\mathrm{Rep}_D((f+g)(\vec{v})) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 7v_1 + 9v_2 \\ 10v_1 + 10v_2 \end{pmatrix}$$

The particular functions in this example have this action on the representations.

$$\mathrm{Rep}_{D}(f(\vec{v})) = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 2v_1 + v_2 \\ 3v_1 + 4v_2 \end{pmatrix}$$

$$\mathrm{Rep}_{D}(g(\vec{v})) = \begin{pmatrix} 5 & 8 \\ 7 & 6 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 5v_1 + 8v_2 \\ 7v_1 + 6v_2 \end{pmatrix}$$

So we want the matrix with this effect.

$$\mathrm{Rep}_{D}((f+g)(\vec{v})) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 7v_1 + 9v_2 \\ 10v_1 + 10v_2 \end{pmatrix}$$

Thus

$$\mathrm{Rep}_{B,D}(f+g) = \begin{pmatrix} 7 & 9 \\ 10 & 10 \end{pmatrix}$$

The particular functions in this example have this action on the representations.

$$\mathrm{Rep}_D(f(\vec{v})) = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 2v_1 + v_2 \\ 3v_1 + 4v_2 \end{pmatrix}$$

$$\mathrm{Rep}_D(g(\vec{v})) = \begin{pmatrix} 5 & 8 \\ 7 & 6 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 5v_1 + 8v_2 \\ 7v_1 + 6v_2 \end{pmatrix}$$

So we want the matrix with this effect.

$$\mathrm{Rep}_D((f+g)(\vec{v})) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 7v_1 + 9v_2 \\ 10v_1 + 10v_2 \end{pmatrix}$$

Thus

$$\mathrm{Rep}_{B,D}(f+g) = \begin{pmatrix} 7 & 9 \\ 10 & 10 \end{pmatrix}$$

and at least in this example the representation of the sum function is the entry-by-entry sum of the representations of the two functions.

Definition of matrix sum and scalar multiple

1.3 *Definition* The *scalar multiple* of a matrix is the result of entry-by-entry scalar multiplication. The *sum* of two same-sized matrices is their entry-by-entry sum.

Definition of matrix sum and scalar multiple

1.3 *Definition* The *scalar multiple* of a matrix is the result of entry-by-entry scalar multiplication. The *sum* of two same-sized matrices is their entry-by-entry sum.

Example Where

$$A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 0 & 2 \\ 9 & -1/2 & 5 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 \\ 8 & -1 \end{pmatrix}$$

Then

$$A + C = \begin{pmatrix} 2 & -1 \\ 10 & 2 \end{pmatrix} \qquad 5B = \begin{pmatrix} 0 & 0 & 10 \\ 45 & -5/2 & 25 \end{pmatrix}$$

Because the sizes don't match, none of these is defined: $A + B$, $B + A$, $B + C$, $C + B$.

1.4 *Theorem* Let $h, g\colon V \to W$ be linear maps represented with respect to bases B, D by the matrices H and G and let r be a scalar. Then with respect to B, D the map $r \cdot h\colon V \to W$ is represented by rH and the map $h + g\colon V \to W$ is represented by $H + G$.

1.4 *Theorem* Let $h, g\colon V \to W$ be linear maps represented with respect to bases B, D by the matrices H and G and let r be a scalar. Then with respect to B, D the map $r \cdot h\colon V \to W$ is represented by rH and the map $h + g\colon V \to W$ is represented by $H + G$.

Proof We can generalize the two examples that started this section. See Exercise 9 . QED

1.4 *Theorem* Let $h, g\colon V \to W$ be linear maps represented with respect to bases B, D by the matrices H and G and let r be a scalar. Then with respect to B, D the map $r \cdot h\colon V \to W$ is represented by rH and the map $h + g\colon V \to W$ is represented by $H + G$.

Proof We can generalize the two examples that started this section. See Exercise 9 . QED

1.6 *Definition* A *zero matrix* has all entries 0. We write $Z_{n\times m}$ or simply Z (another common notation is $0_{n\times m}$ or just 0).

1.4 *Theorem* Let $h, g\colon V \to W$ be linear maps represented with respect to bases B, D by the matrices H and G and let r be a scalar. Then with respect to B, D the map $r \cdot h\colon V \to W$ is represented by rH and the map $h + g\colon V \to W$ is represented by $H + G$.

Proof We can generalize the two examples that started this section. See Exercise 9 . QED

1.6 *Definition* A *zero matrix* has all entries 0. We write $Z_{n\times m}$ or simply Z (another common notation is $0_{n\times m}$ or just 0).

Example The zero matrix is the identity element for matrix addition.

$$\begin{pmatrix} 3 & 1 & 2 \\ 5 & 0 & 9 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 2 \\ 5 & 0 & 9 \end{pmatrix}$$

This reflects the fact that a zero matrix represents a zero function $Z\colon V \to W$, which is the identity elements for function addition.

Matrix Multiplication

Besides scalar multiplication and addition, another operation that we can perform on functions is composition.

Besides scalar multiplication and addition, another operation that we can perform on functions is composition.

2.1 *Lemma* The composition of linear maps is linear.

Besides scalar multiplication and addition, another operation that we can perform on functions is composition.

2.1 *Lemma* The composition of linear maps is linear.

Proof Let $h\colon V \to W$ and $g\colon W \to U$ be linear. The calculation

$$\begin{aligned}g\circ h\left(c_1\cdot\vec{v}_1+c_2\cdot\vec{v}_2\right) &= g\big(\,h(c_1\cdot\vec{v}_1+c_2\cdot\vec{v}_2)\,\big) = g\big(\,c_1\cdot h(\vec{v}_1)+c_2\cdot h(\vec{v}_2)\,\big)\\ &= c_1\cdot g\big(h(\vec{v}_1)\big)+c_2\cdot g\big(h(\vec{v}_2)\big) = c_1\cdot(g\circ h)(\vec{v}_1)+c_2\cdot(g\circ h)(\vec{v}_2)\end{aligned}$$

shows that $g\circ h\colon V \to U$ preserves linear combinations, and so is linear. QED

Besides scalar multiplication and addition, another operation that we can perform on functions is composition.

2.1 *Lemma* The composition of linear maps is linear.

Proof Let $h\colon V \to W$ and $g\colon W \to U$ be linear. The calculation

$$\begin{aligned} g\circ h\,\bigl(c_1\cdot\vec{v}_1+c_2\cdot\vec{v}_2\bigr) &= g\bigl(\,h(c_1\cdot\vec{v}_1+c_2\cdot\vec{v}_2)\,\bigr) = g\bigl(\,c_1\cdot h(\vec{v}_1)+c_2\cdot h(\vec{v}_2)\,\bigr) \\ &= c_1\cdot g\bigl(h(\vec{v}_1)\bigr)+c_2\cdot g\bigl(h(\vec{v}_2)\bigr) = c_1\cdot(g\circ h)(\vec{v}_1)+c_2\cdot(g\circ h)(\vec{v}_2) \end{aligned}$$

shows that $g\circ h\colon V\to U$ preserves linear combinations, and so is linear. QED

We will see how the matrix representations of the two functions combine to make the matrix representation of their composition.

Example Consider two linear functions $h\colon V \to W$ and $g\colon W \to X$ represented as here.

$$\text{Rep}_{B,C}(h) = \begin{pmatrix} 3 & 1 \\ 2 & 5 \\ 4 & 6 \end{pmatrix} \qquad \text{Rep}_{C,D}(g) = \begin{pmatrix} 8 & 7 & 11 \\ 9 & 10 & 12 \end{pmatrix}$$

(The sizes of the matrices show that V has dimension 2, W has dimension 3, and X has dimension 2.) We want to see how these two matrices combine to represent the map $g \circ h\colon V \to X$.

Example Consider two linear functions $h\colon V \to W$ and $g\colon W \to X$ represented as here.

$$\mathrm{Rep}_{B,C}(h) = \begin{pmatrix} 3 & 1 \\ 2 & 5 \\ 4 & 6 \end{pmatrix} \qquad \mathrm{Rep}_{C,D}(g) = \begin{pmatrix} 8 & 7 & 11 \\ 9 & 10 & 12 \end{pmatrix}$$

(The sizes of the matrices show that V has dimension 2, W has dimension 3, and X has dimension 2.) We want to see how these two matrices combine to represent the map $g \circ h\colon V \to X$.

We will compute the action of $g \circ h$ by first computing the action of h on a vector $\vec{v} \in V$.

$$\mathrm{Rep}_C(h(\vec{v})) = \mathrm{Rep}_{B,C}(h) \cdot \mathrm{Rep}_B(\vec{v}) = \begin{pmatrix} 3 & 1 \\ 2 & 5 \\ 4 & 6 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 3v_1 + v_2 \\ 2v_1 + 5v_2 \\ 4v_1 + 6v_2 \end{pmatrix}$$

Apply g to that.

$$\mathrm{Rep}_{C,D}(g)\cdot\mathrm{Rep}_C(h(\vec{v}))=\begin{pmatrix}8&7&11\\9&10&12\end{pmatrix}\begin{pmatrix}3v_1+v_2\\2v_1+5v_2\\4v_1+6v_2\end{pmatrix}$$

$$=\begin{pmatrix}8(3v_1+v_2)+7(2v_1+5v_2)+11(4v_1+6v_2)\\9(3v_1+v_2)+10(2v_1+5v_2)+12(4v_1+6v_2)\end{pmatrix}$$

Gather terms.

$$=\begin{pmatrix}(8\cdot3+7\cdot2+11\cdot4)v_1+(8\cdot1+7\cdot5+11\cdot6)v_2\\(9\cdot3+10\cdot2+12\cdot4)v_1+(9\cdot1+10\cdot5+12\cdot6)v_2\end{pmatrix}$$

Rewrite that as a matrix-vector multiplication.

$$\begin{pmatrix}8\cdot3+7\cdot2+11\cdot4&8\cdot1+7\cdot5+11\cdot6\\9\cdot3+10\cdot2+12\cdot4&9\cdot1+10\cdot5+12\cdot6\end{pmatrix}\begin{pmatrix}v_1\\v_2\end{pmatrix}$$

Here is the combination.

$$\begin{pmatrix}8&7&11\\9&10&12\end{pmatrix}\begin{pmatrix}3&1\\2&5\\4&6\end{pmatrix}=\begin{pmatrix}8\cdot3+7\cdot2+11\cdot4&8\cdot1+7\cdot5+11\cdot6\\9\cdot3+10\cdot2+12\cdot4&9\cdot1+10\cdot5+12\cdot6\end{pmatrix}$$

Definition of matrix multiplication

2.3 *Definition* The *matrix-multiplicative product* of the $m \times r$ matrix G and the $r \times n$ matrix H is the $m \times n$ matrix P, where

$$p_{i,j} = g_{i,1}h_{1,j} + g_{i,2}h_{2,j} + \cdots + g_{i,r}h_{r,j}$$

that is, where the i,j-th entry of the product is the dot product of the i-th row of the first matrix with the j-th column of the second.

$$GH = \begin{pmatrix} & \vdots & & \\ g_{i,1} & g_{i,2} & \ldots & g_{i,r} \\ & \vdots & & \end{pmatrix} \begin{pmatrix} & h_{1,j} & \\ \ldots & h_{2,j} & \ldots \\ & \vdots & \\ & h_{r,j} & \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \ldots & p_{i,j} & \ldots \\ & \vdots & \end{pmatrix}$$

Example

$$\begin{pmatrix} 3 & 1 & 6 \\ 2 & 5 & 9 \end{pmatrix} \begin{pmatrix} 2 & 0 & 4 \\ 1 & -3 & 5 \\ 4 & 2 & 7 \end{pmatrix} = \begin{pmatrix} 31 & 9 & 59 \\ 45 & 3 & 96 \end{pmatrix}$$

Example This product is not defined.

$$\begin{pmatrix} 1 & 3 & -1 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 5 & 7 & 1 \\ 2 & 2 & 0 \end{pmatrix}$$

Example This product is not defined.

$$\begin{pmatrix} 1 & 3 & -1 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 5 & 7 & 1 \\ 2 & 2 & 0 \end{pmatrix}$$

Example Square matrices of the same size have a defined product.

$$\begin{pmatrix} 1 & 3 & -1 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 5 & 7 & 1 \\ 2 & 2 & 0 \\ 1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 10 & 14 & -1 \\ 0 & 0 & 0 \\ 10 & 14 & 2 \end{pmatrix}$$

This reflects the fact that we can compose two functions from a space to itself $f, g \colon V \to V$.

Matrix multiplication represents composition

2.6 *Theorem* A composition of linear maps is represented by the matrix product of the representatives.

Matrix multiplication represents composition

2.6 *Theorem* A composition of linear maps is represented by the matrix product of the representatives.

Proof Let $h\colon V\to W$ and $g\colon W\to X$ be represented by H and G with respect to bases $B\subset V$, $C\subset W$, and $D\subset X$, of sizes n, r, and m. For any $\vec{v}\in V$ the k-th component of $\operatorname{Rep}_C(\,h(\vec{v})\,)$ is

$$h_{k,1}v_1+\cdots+h_{k,n}v_n$$

and so the i-th component of $\operatorname{Rep}_D(\,g\circ h\,(\vec{v})\,)$ is this.

$$\begin{aligned}g_{i,1}\cdot(h_{1,1}v_1+\cdots+h_{1,n}v_n)+g_{i,2}\cdot(h_{2,1}v_1+\cdots+h_{2,n}v_n)&\\ +\cdots+g_{i,r}\cdot(h_{r,1}v_1+\cdots+h_{r,n}v_n)&\end{aligned}$$

Distribute and regroup on the v's.

$$\begin{aligned}&=(g_{i,1}h_{1,1}+g_{i,2}h_{2,1}+\cdots+g_{i,r}h_{r,1})\cdot v_1\\ &\qquad\qquad+\cdots+(g_{i,1}h_{1,n}+g_{i,2}h_{2,n}+\cdots+g_{i,r}h_{r,n})\cdot v_n\end{aligned}$$

Finish by recognizing that the coefficient of each v_j

$$g_{i,1}h_{1,j} + g_{i,2}h_{2,j} + \cdots + g_{i,r}h_{r,j}$$

matches the definition of the i, j entry of the product GH. QED

Arrow diagrams

This pictures the relationship between maps and matrices.

$$
\begin{tikzcd}
 & W_{wrt\ C} \arrow[dr, "g", "G"'] & \\
V_{wrt\ B} \arrow[ur, "h", "H"'] \arrow[rr, "g \circ h", "GH"'] & & X_{wrt\ D}
\end{tikzcd}
$$

Arrow diagrams

This pictures the relationship between maps and matrices.

$$\begin{tikzcd} & W_{wrt\ \mathrm{C}} \arrow[dr, "g", "G"'] & \\ V_{wrt\ \mathrm{B}} \arrow[ur, "h", "H"'] \arrow[rr, "g \circ h", "GH"'] & & X_{wrt\ \mathrm{D}} \end{tikzcd}$$

Above the arrows, the maps show that the two ways of going from V to X, straight over via the composition or else in two steps by way of W, have the same effect

$$\vec{v} \stackrel{g \circ h}{\longmapsto} g(h(\vec{v})) \qquad \vec{v} \stackrel{h}{\longmapsto} h(\vec{v}) \stackrel{g}{\longmapsto} g(h(\vec{v}))$$

(this is just the definition of composition). Below the arrows, the matrices indicate that multiplying GH into the column vector $\mathrm{Rep}_{B}(\vec{v})$ has the same effect as multiplying the column vector first by H and then multiplying the result by G.

$$\mathrm{Rep}_{B,D}(g \circ h) = GH \qquad \mathrm{Rep}_{C,D}(g)\ \mathrm{Rep}_{B,C}(h) = GH$$

Order, dimensions, and sizes

Consider the composition of two function.

First consider the order in which we write the composition. In $g \circ h$ the function written first, g, is the function applied second.

$$\vec{v} \xrightarrow{h} h(\vec{v}) \xrightarrow{g} g(h(\vec{v}))$$

We write g first to match the definition $g \circ h(\vec{v}) = g(\, h(\vec{v})\,)$. That order carries over to matrices: $g \circ h$ is represented by GH.

Order, dimensions, and sizes

Consider the composition of two function.

First consider the order in which we write the composition. In $g \circ h$ the function written first, g, is the function applied second.

$$\vec{v} \overset{h}{\longmapsto} h(\vec{v}) \overset{g}{\longmapsto} g(h(\vec{v}))$$

We write g first to match the definition $g \circ h(\vec{v}) = g(\,h(\vec{v})\,)$. That order carries over to matrices: $g \circ h$ is represented by GH.

Now consider the dimensions of the domain and codomain spaces.

$$\text{dimension } n \text{ space} \xrightarrow{h} \text{dimension } r \text{ space} \xrightarrow{g} \text{dimension } m \text{ space}$$

The representing $m \times n$ matrix GH is the product of an $m \times r$ matrix G and a $r \times n$ matrix H. Briefly, '$m \times r$ times $r \times n$ equals $m \times n$'.

Matrix multiplication is not commutative

Function composition is not a commutative operation — $\cos(x^2)$ is different than $\cos^2(x)$. This also holds in the special case of composition of linear functions.

Matrix multiplication is not commutative

Function composition is not a commutative operation — $\cos(x^2)$ is different than $\cos^2(x)$. This also holds in the special case of composition of linear functions.

Example Changing the order in which we multiply these matrices

$$\begin{pmatrix} 3 & 3 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} -2 & 6 \\ 6 & 5 \end{pmatrix} = \begin{pmatrix} 12 & 33 \\ 24 & 20 \end{pmatrix}$$

changes the result.

$$\begin{pmatrix} -2 & 6 \\ 6 & 5 \end{pmatrix} \begin{pmatrix} 3 & 3 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} -6 & 18 \\ 18 & 38 \end{pmatrix}$$

Matrix multiplication is not commutative

Function composition is not a commutative operation — $\cos(x^2)$ is different than $\cos^2(x)$. This also holds in the special case of composition of linear functions.

Example Changing the order in which we multiply these matrices

$$\begin{pmatrix} 3 & 3 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} -2 & 6 \\ 6 & 5 \end{pmatrix} = \begin{pmatrix} 12 & 33 \\ 24 & 20 \end{pmatrix}$$

changes the result.

$$\begin{pmatrix} -2 & 6 \\ 6 & 5 \end{pmatrix} \begin{pmatrix} 3 & 3 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} -6 & 18 \\ 18 & 38 \end{pmatrix}$$

Example The product of these matrices

$$\begin{pmatrix} 3 & 4 \\ 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} 8 & 12 & 0 \\ -4 & 0 & 1/2 \end{pmatrix}$$

is defined in one order and not defined in the other.

2.12 *Theorem* If F, G, and H are matrices, and the matrix products are defined, then the product is associative $(FG)H = F(GH)$ and distributes over matrix addition $F(G + H) = FG + FH$ and $(G + H)F = GF + HF$.

2.12 *Theorem* If F, G, and H are matrices, and the matrix products are defined, then the product is associative $(FG)H = F(GH)$ and distributes over matrix addition $F(G + H) = FG + FH$ and $(G + H)F = GF + HF$.

Proof Associativity holds because matrix multiplication represents function composition, which is associative: the maps $(f \circ g) \circ h$ and $f \circ (g \circ h)$ are equal as both send $\vec{v}$ to $f(g(h(\vec{v})))$.

2.12 *Theorem* If F, G, and H are matrices, and the matrix products are defined, then the product is associative $(FG)H = F(GH)$ and distributes over matrix addition $F(G + H) = FG + FH$ and $(G + H)F = GF + HF$.

Proof Associativity holds because matrix multiplication represents function composition, which is associative: the maps $(f \circ g) \circ h$ and $f \circ (g \circ h)$ are equal as both send $\vec{v}$ to $f(g(h(\vec{v})))$.

Distributivity is similar. For instance, the first one goes $f \circ (g + h)\,(\vec{v}) = f\big(\,(g + h)(\vec{v})\,\big) = f\big(\,g(\vec{v}) + h(\vec{v})\,\big) = f(g(\vec{v})) + f(h(\vec{v})) = f \circ g(\vec{v}) + f \circ h(\vec{v})$ (the third equality uses the linearity of f). Right-distributivity goes the same way. QED

Mechanics of Matrix Multiplication

Combinatorics of multiplication

The striking thing about matrix multiplication is the way rows and columns combine. Here a second row and a third column combine to make a $2, 3$ entry.

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 4 & 6 & 8 \\ 5 & 7 & 9 \end{pmatrix} = \begin{pmatrix} 9 & 13 & 17 \\ 5 & 7 & 9 \\ 4 & 6 & 8 \end{pmatrix}$$

Combinatorics of multiplication

The striking thing about matrix multiplication is the way rows and columns combine. Here a second row and a third column combine to make a $2,3$ entry.

$$\begin{pmatrix} 1 & 1 \\ \boxed{0 \quad 1} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 4 & 6 & \boxed{8} \\ 5 & 7 & \boxed{9} \end{pmatrix} = \begin{pmatrix} 9 & 13 & 17 \\ 5 & 7 & \boxed{9} \\ 4 & 6 & 8 \end{pmatrix}$$

The i,j entry of the matrix product GH is the dot product of row i of the left matrix G with column j of the right one H.

$$p_{i,j} = g_{i,\boxed{1}} h_{\boxed{1},j} + g_{i,\boxed{2}} h_{\boxed{2},j} + \cdots + g_{i,\boxed{r}} h_{\boxed{r},j}$$

Combinatorics of multiplication

The striking thing about matrix multiplication is the way rows and columns combine. Here a second row and a third column combine to make a $2,3$ entry.

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 4 & 6 & 8 \\ 5 & 7 & 9 \end{pmatrix} = \begin{pmatrix} 9 & 13 & 17 \\ 5 & 7 & 9 \\ 4 & 6 & 8 \end{pmatrix}$$

The i,j entry of the matrix product GH is the dot product of row i of the left matrix G with column j of the right one H.

$$p_{i,j} = g_{i,1}h_{1,j} + g_{i,2}h_{2,j} + \cdots + g_{i,r}h_{r,j}$$

We can view this as the left matrix acting by multiplying its rows into the columns of the right matrix. Or we could see it as the right matrix using its columns to act on the left matrix's rows.

3.2 *Definition* A matrix with all 0's except for a 1 in the i, j entry is an i, j *unit* matrix.

Example The $2, 1$ unit 2×3 matrix multiplies from the left

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 3 \\ 5 & 6 & 4 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

to copy row 1 of the multiplicand into row 2 of the result.

3.2 *Definition* A matrix with all 0's except for a 1 in the i,j entry is an i,j *unit* matrix.

Example The $2,1$ unit 2×3 matrix multiplies from the left

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 3 \\ 5 & 6 & 4 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

to copy row 1 of the multiplicand into row 2 of the result.

Example From the right the $2,1$ unit 2×3 matrix

$$\begin{pmatrix} 3 & 4 \\ 6 & 5 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 \\ 5 & 0 & 0 \end{pmatrix}$$

copies column 2 of the first matrix into column 1 of the result.

3.2 *Definition* A matrix with all 0's except for a 1 in the i,j entry is an i,j *unit* matrix.

Example The $2,1$ unit 2×3 matrix multiplies from the left

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 3 \\ 5 & 6 & 4 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

to copy row 1 of the multiplicand into row 2 of the result.

Example From the right the $2,1$ unit 2×3 matrix

$$\begin{pmatrix} 3 & 4 \\ 6 & 5 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 \\ 5 & 0 & 0 \end{pmatrix}$$

copies column 2 of the first matrix into column 1 of the result.

Example Rescaling the unit matrix rescales the result.

$$\begin{pmatrix} 3 & 4 \\ 6 & 5 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 3 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 12 & 0 & 0 \\ 15 & 0 & 0 \end{pmatrix}$$

3.7 *Lemma* In a product of two matrices G and H, the columns of GH are formed by taking G times the columns of H

$$G \cdot \left(\begin{array}{c|c|c} \vdots & & \vdots \\ \vec{h}_1 & \cdots & \vec{h}_n \\ \vdots & & \vdots \end{array}\right) = \left(\begin{array}{c|c|c} \vdots & & \vdots \\ G \cdot \vec{h}_1 & \cdots & G \cdot \vec{h}_n \\ \vdots & & \vdots \end{array}\right)$$

and the rows of GH are formed by taking the rows of G times H

$$\left(\begin{array}{c} \cdots \ \vec{g}_1 \ \cdots \\ \hline \vdots \\ \hline \cdots \ \vec{g}_r \ \cdots \end{array}\right) \cdot H = \left(\begin{array}{c} \cdots \ \vec{g}_1 \cdot H \ \cdots \\ \hline \vdots \\ \hline \cdots \ \vec{g}_r \cdot H \ \cdots \end{array}\right)$$

(ignoring the extra parentheses).

Proof We will check that in a product of 2×2 matrices, the rows of the product equal the product of the rows of G with the entire matrix H.

$$\begin{aligned}
\begin{pmatrix} g_{1,1} & g_{1,2} \\ g_{2,1} & g_{2,2} \end{pmatrix}\begin{pmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \end{pmatrix} &= \begin{pmatrix} (g_{1,1} \;\; g_{1,2})H \\ (g_{2,1} \;\; g_{2,2})H \end{pmatrix} \\
&= \begin{pmatrix} (g_{1,1}h_{1,1}+g_{1,2}h_{2,1} \;\; g_{1,1}h_{1,2}+g_{1,2}h_{2,2}) \\ (g_{2,1}h_{1,1}+g_{2,2}h_{2,1} \;\; g_{2,1}h_{1,2}+g_{2,2}h_{2,2}) \end{pmatrix}
\end{aligned}$$

We leave the more general check as an exercise. QED

3.8 *Definition* The *main diagonal* (or *principle diagonal* or *diagonal*) of a square matrix goes from the upper left to the lower right.

3.8 *Definition* The *main diagonal* (or *principle diagonal* or *diagonal*) of a square matrix goes from the upper left to the lower right.

3.9 *Definition* An *identity matrix* is square and every entry is 0 except for 1's in the main diagonal.

$$I_{n\times n} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ & \vdots & & \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

3.8 *Definition* The *main diagonal* (or *principle diagonal* or *diagonal*) of a square matrix goes from the upper left to the lower right.

3.9 *Definition* An *identity matrix* is square and every entry is 0 except for 1's in the main diagonal.

$$I_{n\times n} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ & \vdots & & \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

Taking the product with an identity matrix returns the multiplicand.

Example Multiplication by an identity from the left

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ -1 & 5 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -1 & 5 \end{pmatrix}$$

or from the right leaves the matrix unchanged.

$$\begin{pmatrix} 3 & 2 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -1 & 5 \end{pmatrix}$$

3.12 *Definition* A *diagonal matrix* is square and has 0's off the main diagonal.

$$\begin{pmatrix} a_{1,1} & 0 & \ldots & 0 \\ 0 & a_{2,2} & \ldots & 0 \\ & \vdots & & \\ 0 & 0 & \ldots & a_{n,n} \end{pmatrix}$$

3.12 *Definition* A *diagonal matrix* is square and has 0's off the main diagonal.

$$\begin{pmatrix} a_{1,1} & 0 & \ldots & 0 \\ 0 & a_{2,2} & \ldots & 0 \\ & \vdots & & \\ 0 & 0 & \ldots & a_{n,n} \end{pmatrix}$$

Example Multiplication from the left by a diagonal matrix rescales the rows.

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ -1 & 5 \end{pmatrix} = \begin{pmatrix} 6 & 4 \\ -3 & 15 \end{pmatrix}$$

From the right it rescales the columns.

$$\begin{pmatrix} 3 & 2 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 6 & 6 \\ -2 & 15 \end{pmatrix}$$

3.14 *Definition* A *permutation matrix* is square and is all 0's except for a single 1 in each row and column.

3.14 *Definition* A *permutation matrix* is square and is all 0's except for a single 1 in each row and column.

Example Multiplication by a permutation matrix from the left will swap rows.

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{pmatrix}$$

From the right it swaps columns.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 3 \\ 5 & 4 & 6 \\ 8 & 7 & 9 \end{pmatrix}$$

3.19 *Definition* The *elementary reduction matrices* result from applying a one Gaussian operation to an identity matrix.

1) $I \xrightarrow{k\rho_i} M_i(k)$ for $k \neq 0$
2) $I \xrightarrow{\rho_i \leftrightarrow \rho_j} P_{i,j}$ for $i \neq j$
3) $I \xrightarrow{k\rho_i + \rho_j} C_{i,j}(k)$ for $i \neq j$

3.19 *Definition* The *elementary reduction matrices* result from applying a one Gaussian operation to an identity matrix.

1) $I \xrightarrow{k\rho_i} M_i(k)$ for $k \neq 0$

2) $I \xrightarrow{\rho_i \leftrightarrow \rho_j} P_{i,j}$ for $i \neq j$

3) $I \xrightarrow{k\rho_i + \rho_j} C_{i,j}(k)$ for $i \neq j$

Example Here are some 2×2 examples.

$$M_2(3) = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \quad P_{1,2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad C_{1,2}(-3) = \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}$$

3.19 *Definition* The *elementary reduction matrices* result from applying a one Gaussian operation to an identity matrix.

1) $I \xrightarrow{k\rho_i} M_i(k)$ for $k \neq 0$

2) $I \xrightarrow{\rho_i \leftrightarrow \rho_j} P_{i,j}$ for $i \neq j$

3) $I \xrightarrow{k\rho_i + \rho_j} C_{i,j}(k)$ for $i \neq j$

Example Here are some 2×2 examples.

$$M_2(3) = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \quad P_{1,2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad C_{1,2}(-3) = \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}$$

Example Some 3×3 examples.

$$P_{2,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad C_{2,3}(-4) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 0 \end{pmatrix}$$

Example Multiplying on the left by the 3×3 matrix $M_2(1/2)$ has the effect of the row operation $(1/2)\rho_2$.

$$\begin{pmatrix}1&0&0\\0&1/2&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&3&0\\0&2&5\\0&0&0\end{pmatrix}=\begin{pmatrix}1&3&0\\0&1&5/2\\0&0&0\end{pmatrix}$$

Example Multiplying on the left by the 3×3 matrix $M_2(1/2)$ has the effect of the row operation $(1/2)\rho_2$.

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 1/2 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 3 & 0\\ 0 & 2 & 5\\ 0 & 0 & 0\end{pmatrix}=\begin{pmatrix}1 & 3 & 0\\ 0 & 1 & 5/2\\ 0 & 0 & 0\end{pmatrix}$$

Example Left multiplication by $C_{1,3}(-2)$ performs the row operation $-2\rho_1+\rho_3$.

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -2 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 3 & 0\\ 0 & 2 & 5\\ 2 & 1 & 1\end{pmatrix}=\begin{pmatrix}1 & 3 & 0\\ 0 & 2 & 5\\ 0 & -5 & 1\end{pmatrix}$$

3.20 *Lemma* Matrix multiplication can do Gaussian reduction.

1) If $H \xrightarrow{k\rho_i} G$ then $M_i(k)H = G$.

2) If $H \xrightarrow{\rho_i \leftrightarrow \rho_j} G$ then $P_{i,j}H = G$.

3) If $H \xrightarrow{k\rho_i + \rho_j} G$ then $C_{i,j}(k)H = G$.

Proof Clear. QED

3.23 *Corollary* For any matrix H there are elementary reduction matrices $R_1, \ldots, R_r$ such that $R_r \cdot R_{r-1} \cdots R_1 \cdot H$ is in reduced echelon form.

3.20 *Lemma* Matrix multiplication can do Gaussian reduction.

1) If $H \xrightarrow{k\rho_i} G$ then $M_i(k)H = G$.

2) If $H \xrightarrow{\rho_i \leftrightarrow \rho_j} G$ then $P_{i,j}H = G$.

3) If $H \xrightarrow{k\rho_i + \rho_j} G$ then $C_{i,j}(k)H = G$.

Proof Clear. QED

3.23 *Corollary* For any matrix H there are elementary reduction matrices $R_1, \ldots, R_r$ such that $R_r \cdot R_{r-1} \cdots R_1 \cdot H$ is in reduced echelon form.

Example We can bring this augmented matrix to echelon form with matrix multiplication.

$$\left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 2 & -2 & -1 & 6 \\ 0 & 3 & 1 & 5 \end{array}\right)$$

We first perform $-2\rho_1 + \rho_2$ via left multiplication by $C_{1,2}(-2)$.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 2 & -2 & -1 & 6 \\ 0 & 3 & 1 & 5 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & 0 & -5 & -2 \\ 0 & 3 & 1 & 5 \end{array}\right)$$

We first perform $-2\rho_1 + \rho_2$ via left multiplication by $C_{1,2}(-2)$.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 2 & -2 & -1 & 6 \\ 0 & 3 & 1 & 5 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & 0 & -5 & -2 \\ 0 & 3 & 1 & 5 \end{array}\right)$$

Now we swap rows 2 and 3 with $P_{2,3}$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & 0 & -5 & -2 \\ 0 & 3 & 1 & 5 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & 3 & 1 & 5 \\ 0 & 0 & -5 & -2 \end{array}\right)$$

We first perform $-2\rho_1 + \rho_2$ via left multiplication by $C_{1,2}(-2)$.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 2 & -2 & -1 & 6 \\ 0 & 3 & 1 & 5 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & 0 & -5 & -2 \\ 0 & 3 & 1 & 5 \end{array}\right)$$

Now we swap rows 2 and 3 with $P_{2,3}$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & 0 & -5 & -2 \\ 0 & 3 & 1 & 5 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & 3 & 1 & 5 \\ 0 & 0 & -5 & -2 \end{array}\right)$$

When writing them out in full, remember that the matrix used first appears to the right of the matrix used second.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 2 & -2 & -1 & 6 \\ 0 & 3 & 1 & 5 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & 3 & 1 & 5 \\ 0 & 0 & -5 & -2 \end{array}\right)$$

Inverses

Function inverses

We finish this section by considering how to represent the inverse of a linear map.

Function inverses

We finish this section by considering how to represent the inverse of a linear map.

We first recall some things about inverses. Where $\pi\colon \mathbb{R}^3 \to \mathbb{R}^2$ is the projection map and $\iota\colon \mathbb{R}^2 \to \mathbb{R}^3$ is the embedding

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \stackrel{\pi}{\longmapsto} \begin{pmatrix} x \\ y \end{pmatrix} \qquad \begin{pmatrix} x \\ y \end{pmatrix} \stackrel{\iota}{\longmapsto} \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

then the composition $\pi \circ \iota$ is the identity map $\pi \circ \iota = \text{id}$ on $\mathbb{R}^2$.

$$\begin{pmatrix} x \\ y \end{pmatrix} \stackrel{\iota}{\longmapsto} \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \stackrel{\pi}{\longmapsto} \begin{pmatrix} x \\ y \end{pmatrix}$$

We say that ι is a *right inverse* of π or, what is the same thing, that π is a *left inverse* of ι.

However, composition in the other order $\iota \circ \pi$ doesn't give the identity map — here is a vector that is not sent to itself under $\iota \circ \pi$.

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \overset{\pi}{\longmapsto} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \overset{\iota}{\longmapsto} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

However, composition in the other order $\iota \circ \pi$ doesn't give the identity map — here is a vector that is not sent to itself under $\iota \circ \pi$.

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \xmapsto{\pi} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \xmapsto{\iota} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

In fact, π has no left inverse at all. For, if f were to be a left inverse of π then we would have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \xmapsto{\pi} \begin{pmatrix} x \\ y \end{pmatrix} \xmapsto{f} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

for all of the infinitely many z's. But a function f cannot send a single argument $\binom{x}{y}$ to more than one value.

So a function can have a right inverse but no left inverse, or a left inverse but no right inverse. A function can also fail to have an inverse on either side; one example is the zero transformation on $\mathbb{R}^2$.

Some functions have a *two-sided inverse*, another function that is the inverse both from the left and from the right. For instance, the map given by $\vec{v} \mapsto 2 \cdot \vec{v}$ has the two-sided inverse $\vec{v} \mapsto (1/2) \cdot \vec{v}$. The appendix shows that a function has a two-sided inverse if and only if it is both one-to-one and onto. The appendix also shows that if a function f has a two-sided inverse then it is unique, so we call it 'the' inverse and write f^{-1}.

Some functions have a *two-sided inverse*, another function that is the inverse both from the left and from the right. For instance, the map given by $\vec{v} \mapsto 2 \cdot \vec{v}$ has the two-sided inverse $\vec{v} \mapsto (1/2) \cdot \vec{v}$. The appendix shows that a function has a two-sided inverse if and only if it is both one-to-one and onto. The appendix also shows that if a function f has a two-sided inverse then it is unique, so we call it 'the' inverse and write f^{-1}.

In addition, recall that we have shown in Theorem II.2.20 that if a linear map has a two-sided inverse then that inverse is also linear.

Some functions have a *two-sided inverse*, another function that is the inverse both from the left and from the right. For instance, the map given by $\vec{v} \mapsto 2 \cdot \vec{v}$ has the two-sided inverse $\vec{v} \mapsto (1/2) \cdot \vec{v}$. The appendix shows that a function has a two-sided inverse if and only if it is both one-to-one and onto. The appendix also shows that if a function f has a two-sided inverse then it is unique, so we call it 'the' inverse and write f^{-1}.

In addition, recall that we have shown in Theorem II.2.20 that if a linear map has a two-sided inverse then that inverse is also linear.

Thus, our goal in this subsection is, where a linear h has an inverse, to find the relationship between $\mathrm{Rep}_{B,D}(h)$ and $\mathrm{Rep}_{D,B}(h^{-1})$.

Definition of matrix inverse

4.1 *Definition* A matrix G is a *left inverse matrix* of the matrix H if GH is the identity matrix. It is a *right inverse* if HG is the identity. A matrix H with a two-sided inverse is an *invertible matrix*. That two-sided inverse is denoted H^{-1}.

Definition of matrix inverse

4.1 *Definition* A matrix G is a *left inverse matrix* of the matrix H if GH is the identity matrix. It is a *right inverse* if HG is the identity. A matrix H with a two-sided inverse is an *invertible matrix*. That two-sided inverse is denoted H^{-1}.

Example This matrix

$$H = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$$

has a two-sided inverse.

$$H^{-1} = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$$

To check that we can multiply in both orders. Here is one order; the other is just as easy to check.

$$\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

4.2 *Lemma* If a matrix has both a left inverse and a right inverse then the two are equal.

4.2 *Lemma* If a matrix has both a left inverse and a right inverse then the two are equal.

4.3 *Theorem* A matrix is invertible if and only if it is nonsingular.

4.2 *Lemma* If a matrix has both a left inverse and a right inverse then the two are equal.

4.3 *Theorem* A matrix is invertible if and only if it is nonsingular.

Proof *(For both results.)* Given a matrix H, fix spaces of appropriate dimension for the domain and codomain and fix bases for these spaces. With respect to these bases, H represents a map h. The statements are true about the map and therefore they are true about the matrix. QED

4.4 *Lemma* A product of invertible matrices is invertible: if G and H are invertible and GH is defined then GH is invertible and $(GH)^{-1} = H^{-1}G^{-1}$.

4.4 *Lemma* A product of invertible matrices is invertible: if G and H are invertible and GH is defined then GH is invertible and $(GH)^{-1} = H^{-1}G^{-1}$.

Proof Because the two matrices are invertible they are square, and because their product is defined they must both be $n\times n$. Fix bases — the natural choice is the standard basis — to get maps $g, h\colon \mathbb{R}^n \to \mathbb{R}^n$ that are associated with the matrices, $G = \operatorname{Rep}_{\mathcal{E}_n,\mathcal{E}_n}(g)$ and $H = \operatorname{Rep}_{\mathcal{E}_n,\mathcal{E}_n}(h)$.

4.4 *Lemma* A product of invertible matrices is invertible: if G and H are invertible and GH is defined then GH is invertible and $(GH)^{-1} = H^{-1}G^{-1}$.

Proof Because the two matrices are invertible they are square, and because their product is defined they must both be $n \times n$. Fix bases — the natural choice is the standard basis — to get maps $g, h \colon \mathbb{R}^n \to \mathbb{R}^n$ that are associated with the matrices, $G = \mathrm{Rep}_{\mathcal{E}_n,\mathcal{E}_n}(g)$ and $H = \mathrm{Rep}_{\mathcal{E}_n,\mathcal{E}_n}(h)$.

Consider $h^{-1}g^{-1}$. By the prior paragraph this composition is defined. This map is a two-sided inverse of gh since $(h^{-1}g^{-1})(gh) = h^{-1}(\mathrm{id})h = h^{-1}h = \mathrm{id}$ and $(gh)(h^{-1}g^{-1}) = g(\mathrm{id})g^{-1} = gg^{-1} = \mathrm{id}$. The matrices representing the maps reflect this equality. QED

4.4 *Lemma* A product of invertible matrices is invertible: if G and H are invertible and GH is defined then GH is invertible and $(GH)^{-1} = H^{-1}G^{-1}$.

Proof Because the two matrices are invertible they are square, and because their product is defined they must both be $n \times n$. Fix bases — the natural choice is the standard basis — to get maps $g, h \colon \mathbb{R}^n \to \mathbb{R}^n$ that are associated with the matrices, $G = \mathrm{Rep}_{\mathcal{E}_n,\mathcal{E}_n}(g)$ and $H = \mathrm{Rep}_{\mathcal{E}_n,\mathcal{E}_n}(h)$.

Consider $h^{-1}g^{-1}$. By the prior paragraph this composition is defined. This map is a two-sided inverse of gh since $(h^{-1}g^{-1})(gh) = h^{-1}(\mathrm{id})h = h^{-1}h = \mathrm{id}$ and $(gh)(h^{-1}g^{-1}) = g(\mathrm{id})g^{-1} = gg^{-1} = \mathrm{id}$. The matrices representing the maps reflect this equality. QED

Here is the arrow diagram for matrix inverses.

$$
\begin{tikzcd}
 & W_{\mathit{wrt}\ C} \arrow[dr, "h^{-1}", "H^{-1}"'] & \\
V_{\mathit{wrt}\ B} \arrow[ur, "h", "H"'] \arrow[rr, "\mathrm{id}", "I"'] & & V_{\mathit{wrt}\ B}
\end{tikzcd}
$$

4.7 *Lemma* A matrix H is invertible if and only if it can be written as the product of elementary reduction matrices. We can compute the inverse by applying to the identity matrix the same row steps, in the same order, that Gauss-Jordan reduce H.

4.7 *Lemma* A matrix H is invertible if and only if it can be written as the product of elementary reduction matrices. We can compute the inverse by applying to the identity matrix the same row steps, in the same order, that Gauss-Jordan reduce H.

Proof The matrix H is invertible if and only if it is nonsingular and thus Gauss-Jordan reduces to the identity. By Corollary 3.23 we can do this reduction with elementary matrices.

$$R_r \cdot R_{r-1} \ldots R_1 \cdot H = I \qquad (*)$$

4.7 *Lemma* A matrix H is invertible if and only if it can be written as the product of elementary reduction matrices. We can compute the inverse by applying to the identity matrix the same row steps, in the same order, that Gauss-Jordan reduce H.

Proof The matrix H is invertible if and only if it is nonsingular and thus Gauss-Jordan reduces to the identity. By Corollary 3.23 we can do this reduction with elementary matrices.

$$R_r \cdot R_{r-1} \ldots R_1 \cdot H = I \qquad (*)$$

For the first sentence of the result, note that elementary matrices are invertible because elementary row operations are reversible, and that their inverses are also elementary. Apply R_r^{-1} from the left to both sides of $(*)$. Then apply R_{r-1}^{-1}, etc. The result gives H as the product of elementary matrices $H = R_1^{-1} \cdots R_r^{-1} \cdot I$. (The I there covers the case $r = 0$.)

4.7 *Lemma* A matrix H is invertible if and only if it can be written as the product of elementary reduction matrices. We can compute the inverse by applying to the identity matrix the same row steps, in the same order, that Gauss-Jordan reduce H.

Proof The matrix H is invertible if and only if it is nonsingular and thus Gauss-Jordan reduces to the identity. By Corollary 3.23 we can do this reduction with elementary matrices.

$$R_r \cdot R_{r-1} \ldots R_1 \cdot H = I \qquad (*)$$

For the first sentence of the result, note that elementary matrices are invertible because elementary row operations are reversible, and that their inverses are also elementary. Apply R_r^{-1} from the left to both sides of $(*)$. Then apply R_{r-1}^{-1}, etc. The result gives H as the product of elementary matrices $H = R_1^{-1} \cdots R_r^{-1} \cdot I$. (The I there covers the case $r = 0$.)

For the second sentence, group $(*)$ as $(R_r \cdot R_{r-1} \ldots R_1) \cdot H = I$ and recognize what's in the parentheses as the inverse $H^{-1} = R_r \cdot R_{r-1} \ldots R_1 \cdot I$. Restated: applying R_1 to the identity, followed by R_2, etc., yields the inverse of H. QED

Example This matrix is nonsingular and so invertible.

$$A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & -1 \\ 1 & 2 & 0 \end{pmatrix}$$

To ease the calculation called for in the Lemma, we write the matrix A next to the 3×3 identity and as we Gauss-Jordan reduce A we apply those operations to I.

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 2 & 0 & -1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[-\rho_1+\rho_3]{-2\rho_1+\rho_2} \left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & -6 & -3 & -2 & 1 & 0 \\ 0 & -1 & -1 & -1 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{-1/6\rho_2+\rho_3} \left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & -6 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1/2 & -2/3 & -1/6 & 1 \end{array}\right)$$

$$\xrightarrow[-2\rho_3]{-1/6\rho_2} \left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1/2 & 1/3 & -1/6 & 0 \\ 0 & 0 & 1 & 4/3 & 1/3 & -2 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1/2 & 1/3 & -1/6 & 0 \\ 0 & 0 & 1 & 4/3 & 1/3 & -2 \end{array}\right) \xrightarrow[-(1/2)\rho_3+\rho_2]{-\rho_3+\rho_1} \left(\begin{array}{ccc|ccc} 1 & 3 & 0 & -1/3 & -1/3 & 2 \\ 0 & 1 & 0 & -1/3 & -1/3 & 1 \\ 0 & 0 & 1 & 4/3 & 1/3 & -2 \end{array}\right)$$

$$\xrightarrow{-3\rho_2+\rho_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 2/3 & 2/3 & -1 \\ 0 & 1 & 0 & -1/3 & -1/3 & 1 \\ 0 & 0 & 1 & 4/3 & 1/3 & -2 \end{array}\right)$$

We have calculated that this is the inverse.

$$A^{-1} = \begin{pmatrix} 2/3 & 2/3 & -1 \\ -1/3 & -1/3 & 1 \\ 4/3 & 1/3 & -2 \end{pmatrix}$$

Finding the inverse of a matrix A is a lot of work but once we have it then solving linear systems $A\vec{x} = \vec{b}$ is easy.

Example The linear system

$$\begin{aligned} x + 3y + z &= \ \ 2 \\ 2x \qquad\quad - z &= 12 \\ x + 2y \qquad &= \ \ 4 \end{aligned}$$

is this matrix equation.

$$\begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & -1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 12 \\ 4 \end{pmatrix}$$

Solve it by multiplying both sides from the left by the inverse that we found earlier.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2/3 & 2/3 & -1 \\ -1/3 & -1/3 & 1 \\ 4/3 & 1/3 & -2 \end{pmatrix} \begin{pmatrix} 2 \\ 12 \\ 4 \end{pmatrix} = \begin{pmatrix} 16/3 \\ -2/3 \\ -4/3 \end{pmatrix}$$

The inverse of a 2×2 matrix

4.11 *Corollary* The inverse for a 2×2 matrix exists and equals

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

if and only if $ad - bc \neq 0$.

The inverse of a 2×2 matrix

4.11 *Corollary* The inverse for a 2×2 matrix exists and equals

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

if and only if $ad - bc \neq 0$.

Proof This computation is Exercise 21 . QED

The inverse of a 2×2 matrix

4.11 *Corollary* The inverse for a 2×2 matrix exists and equals

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

if and only if $ad - bc \neq 0$.

Proof This computation is Exercise 21 . QED

Example

$$\begin{pmatrix} 2 & 4 \\ -1 & 1 \end{pmatrix}^{-1} = \frac{1}{6}\begin{pmatrix} 1 & -4 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1/6 & -2/3 \\ 1/6 & 1/3 \end{pmatrix}$$

The inverse of a 2×2 matrix

4.11 *Corollary* The inverse for a 2×2 matrix exists and equals

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

if and only if $ad - bc \neq 0$.

Proof This computation is Exercise 21 . QED

Example

$$\begin{pmatrix} 2 & 4 \\ -1 & 1 \end{pmatrix}^{-1} = \frac{1}{6}\begin{pmatrix} 1 & -4 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1/6 & -2/3 \\ 1/6 & 1/3 \end{pmatrix}$$

The 3×3 formula is much more complicated. We will cover it in the next chapter.